Manchester Metropolitan University

e-space

Manchester Metropolitan University's Research Repository

Please cite the Published Version

Baker, F, Smith, CL and Cavan, GM (2018) A Combined Approach to Classifying Land Surface Cover of Urban Domestic Gardens Using Citizen Science Data and High Resolution Image Analysis. Remote Sensing, 10 (4). ISSN 2072-4292

DOI: https://doi.org/10.3390/rs10040537

Publisher: MDPI AG

Downloaded from: https://e-space.mmu.ac.uk/620307/

Usage rights: CC BY Creative Commons: Attribution 4.0

Additional Information: This is an Open Access article published in Remote Sensing, published by MDPI, copyright The Author(s).

Enquiries:
If you have questions about this document, contact openresearch@mmu.ac.uk. Please include the URL of the record in e-space. If you believe that your, or a third party's rights have been compromised through this document please see our Take Down policy (available from https://www.mmu.ac.uk/library/using-the-library/policies-and-guidelines)

remote sensing

MDPI

Article

A Combined Approach to Classifying Land Surface Cover of Urban Domestic Gardens Using Citizen Science Data and High Resolution Image Analysis

Fraser Baker [1], Claire L. Smith [2] and Gina Cavan [1,*]

[1] School of Science and the Environment, Manchester Metropolitan University, John Dalton Building, Chester Street, Manchester M1 5GD, UK; fraser.a.baker@stu.mmu.ac.uk
[2] Department of Geography, University of Leicester, University Road, Leicester LE1 7RH, UK; cls53@le.ac.uk
* Correspondence: g.cavan@mmu.ac.uk; Tel.: +44-161-247-1571

Received: 27 February 2018; Accepted: 28 March 2018; Published: 31 March 2018



Abstract: Domestic gardens are an important component of cities, contributing significantly to urban green infrastructure (GI) and its associated ecosystem services. However, domestic gardens are incredibly heterogeneous which presents challenges for quantifying their GI contribution and associated benefits for sustainable urban development. This study applies an innovative methodology that combines citizen science data with high resolution image analysis to create a garden dataset in the case study city of Manchester, UK. An online Citizen Science Survey (CSS) collected estimates of proportional coverage for 10 garden land surface types from 1031 city residents. High resolution image analysis was conducted to validate the CSS estimates, and to classify 7 land surface cover categories for all garden parcels in the city. Validation of the CSS land surface estimations revealed a mean accuracy of 76.63% (s = 15.24%), demonstrating that citizens are able to provide valid estimates of garden surface coverage proportions. An Object Based Image Analysis (OBIA) classification achieved an estimated overall accuracy of 82%, with further processing required to classify shadow objects. CSS land surface estimations were then extrapolated across the entire classification through calculation of within image class proportions, to provide the proportional coverage of 10 garden land surface types (buildings, hard impervious surfaces, hard pervious surfaces, bare soil, trees, shrubs, mown grass, rough grass, cultivated land, water) within every garden parcel in the city. The final dataset provides a better understanding of the composition of GI in domestic gardens and how this varies across the city. An average garden in Manchester has 50.23% GI, including trees (16.54%), mown grass (14.46%), shrubs (9.19%), cultivated land (7.62%), rough grass (1.97%) and water (0.45%). At the city scale, Manchester has 49.0% GI, and around one fifth (20.94%) of this GI is contained within domestic gardens. This is useful evidence to inform local urban development policies.

Keywords: urban gardens; land surface cover; citizen science; land cover classification; urban green space; green infrastructure; object-based image classification

1. Introduction

Private domestic gardens can significantly contribute to a city's green infrastructure (GI), potentially occupying over a third of a city's urban surface area [1]. While at the individual scale, the influence of a private garden may appear negligible, collectively gardens provide valuable ecosystem services, particularly within densely built up urban environments. The infrastructure and surface cover associated with urban environments renders these areas particularly vulnerable to climate change and extreme meteorological events. These small pockets of urban greenspace, however, perform an important role in the regulation of climate and reduction of associated climate risks such as flooding

and the urban heat island [2]. In spite of this, private domestic gardens have received comparatively little attention within the GI literature [3], most likely due to the small parcel size associated with individual gardens and the lack of regulation, which grants relative freedom for homeowners to adapt and alter garden composition. The size and diversity of private domestic gardens create challenges for quantifying their GI contribution. However, it is important that this aspect of GI is accurately mapped as current approaches can lead to an overestimation of urban greenspace, in addition to subsequent erroneous evaluation of ecosystem service provision and environmental deprivation. This has implications for the future resilience of an urban environment and the health and well-being of its citizens.

The small spatial scale and diversity of gardens implies that local knowledge is a valuable asset for assessing garden composition. Moreover, the impact of a garden on ecosystem functions and services will be greatest within the immediate neighbourhood. The connectivity between garden composition and the surrounding neighbourhood provides rationale for taking a citizen science approach to data collection related to gardens. Further potential benefits of working with local communities to create a garden database include the exchange of knowledge between local people, scientists and stakeholders, enhanced community cohesion and greater transparency for local decision-making [4]. Local surveys have been used previously to gather data on garden use and maintenance, together with socio-economic factors [5,6], however, the role of crowd-sourcing for the accurate characterisation of garden composition has not previously been attempted. There is clear value in working with local residents to collate data on garden GI quantity and quality, and a well-designed citizen science tool has the further advantage that it creates a channel for communication between scientists, decision-makers and the local community. This is particularly important in the context of domestic gardens which are largely unregulated in the UK. Furthermore, the general public is often unaware of the environmental value of their own garden and how they could improve it. Thus, this approach also creates an opportunity to educate individuals and influence behaviour in urban spaces that may previously have fallen outside of the local governance remit, but which have clear environmental consequences for the local neighbourhood [7–10].

The advent of high resolution Earth Observation imagery also provides a viable opportunity to improve estimates of garden composition [1]. Object Based Image Analysis (OBIA) approaches using Ikonos and Worldview-2 imagery have been applied successfully to the garden classification problem, with quoted accuracies of greater than 85% [1,11]. These studies demonstrate the capability for distinguishing between vegetated and non-vegetated areas at the finer spatial scale but lack the detailed information related to GI type, structure and height, which is imperative for rigorous environmental modelling. It is recognised that the functionality of different GI types (e.g., mown grass, rough grass, shrubs, trees, water bodies) is variable and, furthermore, that the ecosystem service capabilities of individual plant species can vary depending on the local environment and the specific service under consideration [12]. Consequently, achieving a classification of gardens which goes beyond a broad vegetated versus non-vegetated classification is necessary for the identification and prioritisation of areas of GI need and it is essential that these fine-scale data are included in multi-scale evaluations.

This article will outline the development of a novel approach to the land surface cover classification within urban domestic gardens which employs a two-stage process. Initially, quantitative data relating to garden composition are captured using an online citizen science survey tool. This dataset is subsequently validated and extended through the classification of high-resolution optical imagery, to produce a high-resolution garden dataset that quantifies the proportion of ten land surface cover types across a post-industrial temperate city. This methodology has the further advantage that the citizen science survey acts as a tool to inform and engage the public about the role of their garden from the perspective of climate regulation and can be used to foster community engagement and resilience.

2. Materials and Methods

2.1. Overview

Manchester, UK, with a total population of 541,263 people (2016 mid-year estimate) and an area of 115.6 km^2, offers an interesting case study for garden investigation and mapping as it has been the focus of significant GI research previously [13–16] and the city is recognised as being especially vulnerable to extreme weather events [17–20]. Manchester comprises 226,640 households of which approximately 68% is private stock (owner occupied and private rented) and the remainder held by Local Authority or Registered Social Landlords (Housing Associations) [21]. Manchester has a disproportionate number of flats/maisonettes/apartments (26.6%) and terraced housing (36.0%) relative to the national housing stock, and detached housing is particularly under-represented (4.3% of Manchester's housing stock) [22]. Gardens comprise over one fifth (20.4%) of the total land area of Manchester, and garden coverage (as a proportion of total land cover) varies considerably across the city, between 0.5% in the City Centre ward to 47% in suburban areas (Figure 1). Domestic gardens in Manchester are therefore a spatially heterogeneous but significant factor in climate-landscape interactions in the area. While assessments of public urban GI and detailed tree assessments have been undertaken previously, little is known about the composition of GI in domestic gardens and how this may vary across the city. A more reliable assessment of garden GI composition and its variability across the city will allow for a more robust framework to inform planning and investment decisions relating to GI solutions within and beyond domestic gardens, and to strengthen the functionality of ecosystems in areas of GI need.

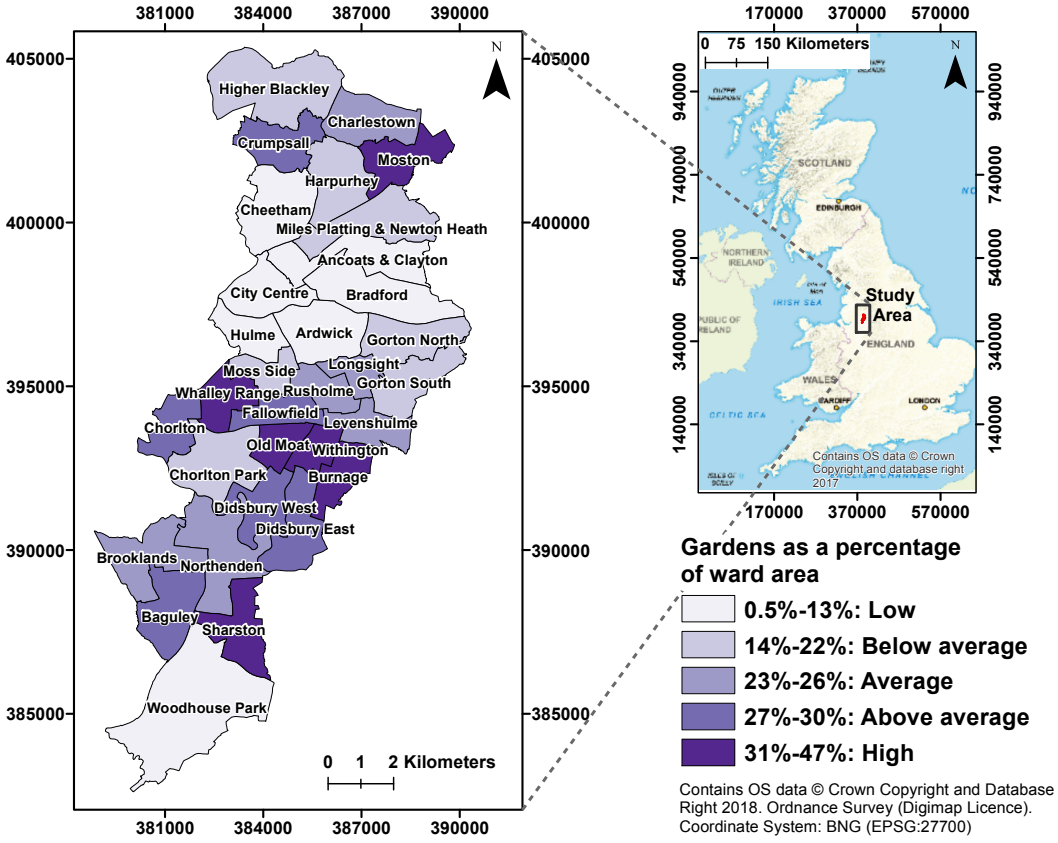


Figure 1. Gardens as a percentage of ward area (groups classified by standard deviation).

2.2. Online Citizen Science Survey Tool: My Back Yard

A short survey designed to be completed by the owner or resident of a household was devised to gather quantitative information about the garden layout and structure. Here, the term garden refers to all of the privately accessible outdoor space associated with a respondent's home address. Such private areas can be geo-located as unique polygon parcels within mapped Ordnance Survey data (i.e., front, back and side garden areas). The survey included specific questions about the proportional cover of ten common garden land surface cover categories (buildings, hard impervious, hard pervious, bare soil, trees, shrubs, mown grass, rough grass, cultivated and water) in order to quantify the amount of green and blue (water) space within an individual domestic garden (Figure 2), and to understand the proportion of other land surfaces, required for subsequent ecosystem service mapping. The Citizen Science Survey (CSS) was hosted online within a tool, which comprised traditional form elements, plus interactive mapping, fulfilling two purposes:

- To gather the proportional land surface cover and exact spatial location of the garden, i.e., the respondent enters their postcode; the tool uses this to determine the boundary of the space-based upon Ordnance Survey mapping; the respondent confirms this or corrects it using spatial editing and selection tools on the map.
- To allow the respondent (and other users) to explore the data already collected, in a generalised form, so that they can see their own contribution in the context of their neighbourhood, and learn about the benefits of green and blue space.

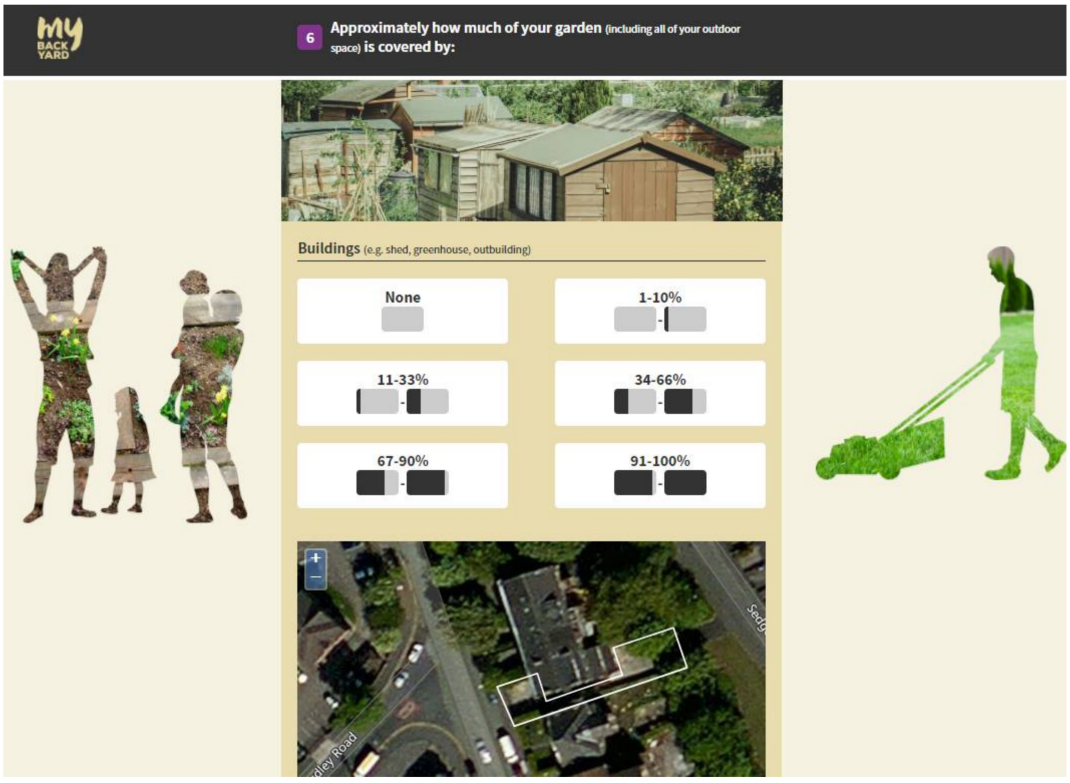


Figure 2. The My Back Yard online survey tool—an example question about the proportional surface cover for buildings.

Minor adjustments to the CSS, which was co-designed with project stakeholders, occurred after further consultation and piloting of the online survey tool. The online survey tool, promoted as *My Back Yard*, was open for a period of 6 months (July 2016–December 2016) for anyone with a UK postcode to complete, however the results presented here are for Manchester, UK, the geographical

focus of this study. Further survey questions, concerning species richness, socio-economic information, garden value, and use, were also included, but presentation and analysis of these is outside the scope of this article.

The *My Back Yard* online survey tool was promoted through a variety of channels including social media (Facebook and Twitter), local press, face-to-face outreach events and via local stakeholders, including the city council, housing and charitable community networks. Respondents were self-selecting and there was potential for a disproportionate response from participants with large, relatively green gardens. In order to achieve a representative sample of gardens from across the case study area, participants in poorly represented wards and housing types were targeted via local community networks. In addition, the online survey tool was promoted as "*My Back Yard*" to encourage participation from those with very small yards and/or little green space in their gardens to complete the survey.

2.3. Extension of the Citizen Science Survey (CSS) Surface Estimations

Image analysis was required to first validate, and then extend, the CSS database, by categorising broad surface cover types for all digitised garden parcels in Manchester. Validation is an important exercise, providing both an estimation of the accuracy of the CSS surface estimations as a whole, and additional information on possible issues with the CSS methodology [23]. Classifying broad surface categories, representing groupings of CSS surfaces for all gardens in the study area, provides a feasible method of extrapolation, whereby the spatial coverage of the general classification is enhanced with detailed information from the CSS surface estimates.

2.3.1. Data

High Resolution True-colour Aerial Imagery (TAI) (spatial scale 0.125 m) was obtained from Getmapping [24]. The TAI data for Manchester was available for June in 2009, 2010, or 2015 as 1 km^2 image tiles, depending upon the specific area of the city, with the most recent (2015) data being available for approximately two thirds of the city (Supplementary Materials, Figure S1). The TAI data enabled the fine-scale identification of surfaces within the urban environment, enabling both manual validation and automatic classification. Pre-processing of TAI was not required as suitable radiometric and georectification calibration was conducted by the data provider.

Vector polygon data representing digital Ordnance Survey garden areas (OSG) were identified from the digital Ordnance Survey Mastermap Topography layer, July 2016 update (OSMT). Further pre-processing was conducted to ensure multiple OSG polygons, representing various areas of a singular garden parcel (e.g., front and back gardens), were united to form singular parcels for further analysis in both the validation and classification exercise. Coverage of OSG for each of the image acquisition dates was 19.57%, 0.19% and 80.24% for the years 2009, 2010 and 2015 respectively. The majority of garden polygons are therefore represented by up to date image data.

Digitised tree canopy data was available from a local environmental NGO: City of Trees [25], for identifying tree canopy areas within garden polygons in the classification exercise (Section 2.3.3) (Figure 3). In addition, building polygons attached to garden areas (Ordnance Survey Buildings: OSB) were extracted from OSMT and classified as either garden buildings or main dwellings for further use in the classification scheme (Section 2.3.3) (Figure 3).

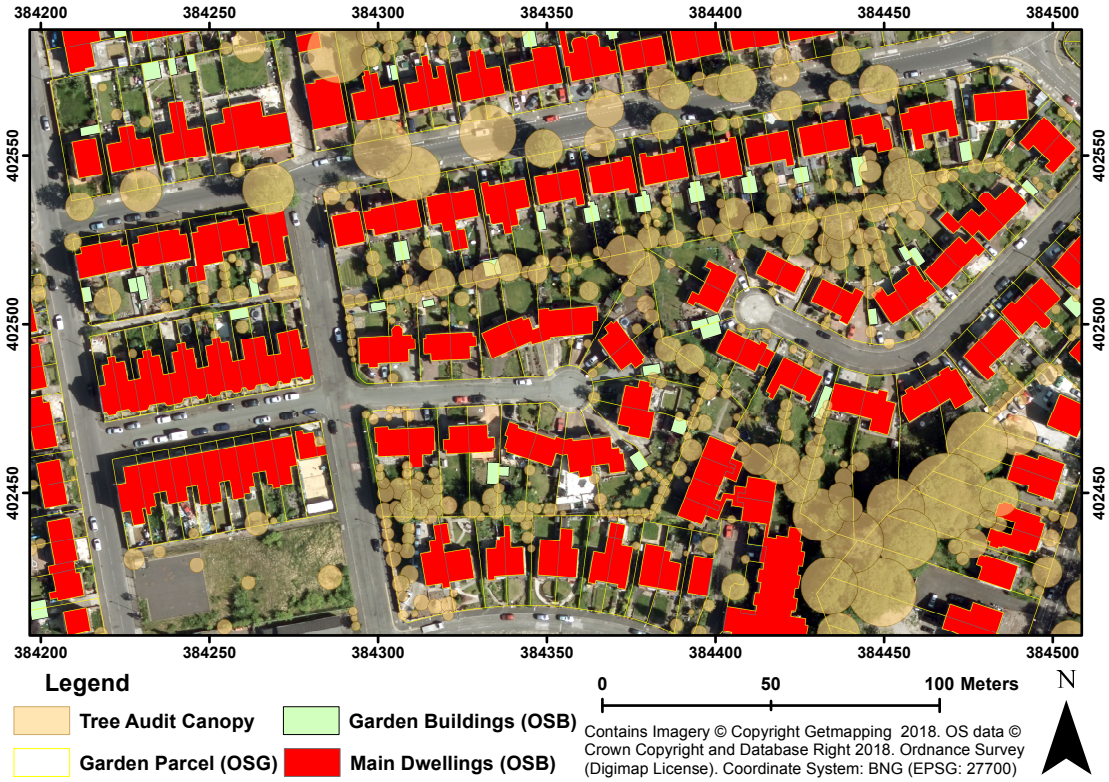


Figure 3. Geospatial datasets used in the image analysis.

2.3.2. Validation of Citizen Science Survey Land Surface Estimations

The garden surface coverage estimations from the CSS were ascertained as fuzzy estimates for each land surface type in one of six estimation categories (None, 1–10%, 11–33%, 34–66%, 67–90% or 91–100%). Respondents were allowed to choose any combination of estimation categories, and while they were told to adjust these if 100% was not within the summed range estimation, respondents could continue the survey regardless of whether the fuzzy estimations tallied to valid estimation. Valid response estimation therefore satisfied the condition:

$$\sum_{l \in Mn} l \leq 100 \leq \sum_{u \in Mx} u \quad (1)$$

where $Mn = \{l_i, \ldots, l_{10}\}$, $Mx = \{u_i, \ldots, u_{10}\}$, i is the index number for a CSS surface, l_i is the lower bracket figure for a surface estimation range, u_i is the upper bracket figure for a surface estimation range. The category *None* therefore has $l = 0$ and $u = 0$. CSS responses that satisfied the above condition were considered useable for further extrapolation.

A manual digitisation method using TAI was devised to assess the accuracy of individual CSS responses (based on [8]). In order to provide an estimation of average Validation Accuracy (VA), a sample of 252 responses were chosen for validation (95% confidence level). The sample was stratified according to garden green infrastructure composition and garden size. For each survey response, the appropriate OSG polygon was geolocated and superimposed onto the image. Digitised polygons were then drawn around homogenous surface areas representing one of five surface groups (Validation Surface Group, VSG) within OSG polygon extents. Multiple CSS surfaces were assigned to a particular class where appropriate, as it proved difficult to distinguish between them in the imagery itself (Table 1). Where image ground features were obscured due to shadow coverage or vertical features (as a result of slightly off-nadir image acquisition), and could not be subsequently interpolated through topological reasoning (e.g., continuation of a hard-surface driveway

into a building) then proportional re-assignment of the obscured area was undertaken and weighted towards under-represented validation classes. For each CSS surface class, hard proportion values were weighted from fuzzy estimations, using the below equations:

$$MN = \sum_{l \in Mn} l \tag{2}$$

$$MX = \sum_{u \in Mn} u \tag{3}$$

$$SP_i = Mn_i + [\,(Mx_i - Mn_i) \cdot \left(\frac{100 - MN}{MX - MN} \right)\,] \tag{4}$$

where MN and MX are the sum of Mn and Mx values, and SP is the calculated hard CSS surface proportion value. SP values were then amalgamated to provide hard values (v) for the relevant validation surface group. Garden proportions calculated from the digitised garden validation polygons were compared to these values to calculate the validation agreement (VA) metric. VA works on the principle that ideal agreement between the hard value CSS validation and digitised validation surface proportions should match exactly. In this instance over-estimation or under-estimation within a validation category has a knock-on effect in other categories. Therefore, validation dis-agreement (VD) was first calculated as the sum of all surface under-estimation (Figure 4).

$$Rc = \{r_i, \ldots, r_5\} \tag{5}$$

where ri = di − vi; i is the reference number for validation surface group, d is the digitised surface proportion and e is the estimated validation surface proportion. Over-estimated surfaces are removed to provide the final metric:

$$Dc = \{\, r \in Rc \mid r > 0 \,\} \tag{6}$$

$$VD = \sum_{r \in Dc} r \tag{7}$$

$$VA = 100 - VD \tag{8}$$

Table 1. Validation Surface Group (VSG) assignment.

Validation Surface Group (VSG)	Associated CSS Surfaces	ASSIGNMENT PROCESS
Buildings	Buildings	Structures identified in image including main parcel dwelling area (if required), outbuildings, extensions and building areas obscuring garden areas
Grass	Mown Grass and Rough Grass	Grass areas in scene
Manmade	Hard Impervious and Hard Pervious	Objects/surfaces representing ground-covering artificial surfaces such as decking, tarmac, concrete, paving slabs (low texture)
Shrubs	Shrubs, Cultivated, Bare Soil and Water	Non grass or tree vegetated areas—includes flower beds, areas of soil (potentially confused with cultivated areas), water (small water bodies are difficult to identify in the imagery)
Trees	Trees	Tall-standing tree canopy areas

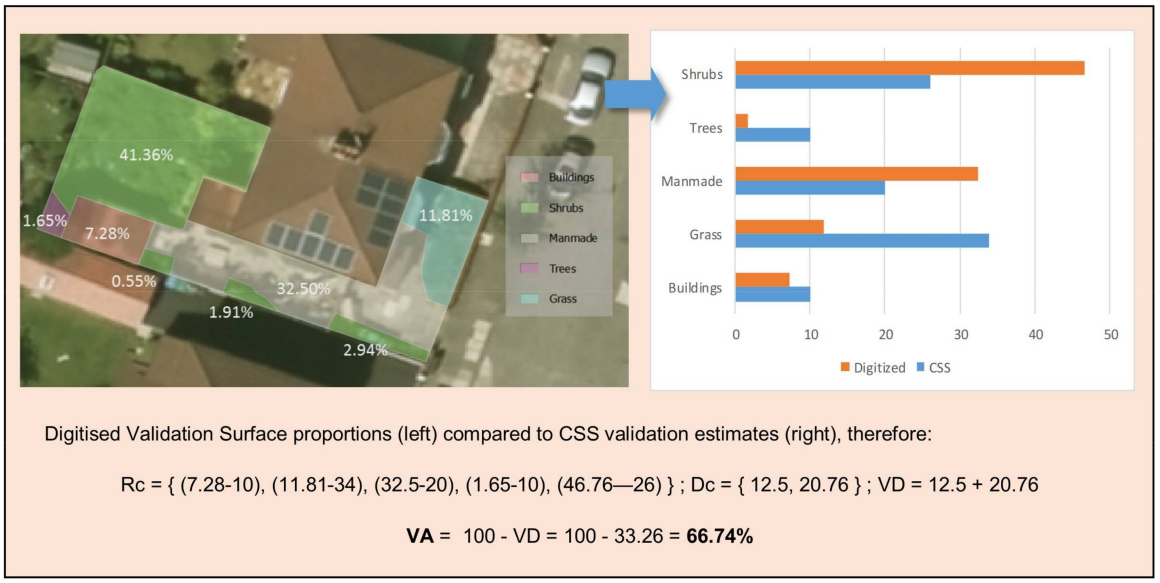


Figure 4. A worked example of the validation methodology.

2.3.3. Classification

Object Based Image Analysis (OBIA) approaches, which are suitable for the classification of high spatial resolution imagery [26], were adopted in this study. OBIA, through image segmentation, enables the creation of a large number of within object features, enhancing spectral information inherent in the scene [27,28]. In addition, as distinct surface areas can form in managed patterns across an urban area, OBIA topological processing enables object classification through defined spatial relationships between object neighbours [29,30]. Such approaches proved useful in the development of the classification ruleset to discriminate between spectrally similar classes. OBIA methods were implemented using the commercial eCognition Developer (Trimble) software.

Initial ruleset experimentation enabled the development of a classification scheme, where class objects were identified as potentially classifiable at acceptable levels of accuracy (approximately above 85%) [31]. Classes were directly related to CSS surface groups, to enable post-classification extrapolation (Table 2). Bare earth was related to the same CSS surface categories as the Shrubs image class, as it was envisaged that CSS respondents may estimate areas of bare soil intermixed with non-grassy vegetation (represented as the shrubs image class) in the CSS. In contrast, areas of open soil, represented as the bare earth image class, may be interpreted as cultivated land or shrubs. Both bare earth and shrubs were thus related to the same CSS surfaces, to enable sufficient extrapolation of bare soil/cultivated/shrubs CSS estimates within image classes containing either surface type.

Table 2. Image classification scheme.

Class Category (Associated Superclass in Brackets)	Description	Related CSS Surface Categories
Bare earth (*non-vegetation*)	Natural exposed non-vegetative surfaces (differentiated from artificial non-vegetative surfaces)	Bare soil, cultivated, shrubs, water
Buildings (*non-vegetation*)	Permanent buildings within garden areas, including garages and sheds, excluding main dwellings	Buildings
Grass (*vegetation*)	Grassed surfaces	Mown grass, rough grass
Manmade (*non-vegetation*)	Permanent non-vegetative manmade surfaces e.g., asphalt drives, decking, gravel, garden furniture	Hard impervious, hard pervious
Shrubs (*vegetation*)	Rough vegetation e.g., shrubs, flower beds, bushes (includes ponds and other water features which are typically covered with aquatic vegetation in the imagery, and are thus spectrally similar to shrubs)	Bare soil, cultivated, shrubs, water

Table 2. *Cont.*

Class Category (Associated Superclass in Brackets)	Description	Related CSS Surface Categories
Trees (*vegetation*)	Tree canopies identified within general confines of tree vector data	Trees
Shadow (*shadow*)	Surfaces completely obscured by shadow	n/a—class reassignment required (Section 2.3.3)

A number of image layers were used to enhance the limited spectral resolution of TAI. Object-based features were typically derived as the mean object value for a particular image layer, with additional measures derived from combinations of image mean features (Table 3). Textural features are used effectively for classification with very high resolution imagery in other studies [32,33], but were found to be too computationally expensive to implement in this study, due to resolution of the data. Thus, a high degree of object feature overlap between the image classes, negated the use of typical machine learning classifiers, in favour of novel iterative topological region growing routines, which take advantage of the wealth of spatial information in the TAI and ancillary data (tree canopy data and OSB polygons) [34]. The development of a fixed ruleset also enabled processing of the TAI on a tile-by-tile basis, thus reducing the extensive processing costs associated with the TAI. The key processes within each module are outlined here. Further detail, including threshold parameters and method diagrams, are provided in the Supplementary Materials for reference.

Table 3. Descriptions of spectrally-derived features.

Feature Layer	Description
Red Green Blue	Normalised prior to creation of additional features. Normalisation for each layer calculated by dividing layer pixel values by the maximum permitted value (in this case 255)
MeanRGB	Simple mean of pixel RGB layer values. Provides approximation of panchromatic data for segmentation, as well as some measure of the general illumination of pixels
SdRGB	Standard Deviation of pixel RGB values. Typical artificial surfaces in the imagery represented by neutral Grey, White and Black. In comparison to more vibrant colours (e.g., representing vegetative surfaces), neutral tones contain a degree of saturation, and have relatively similar values in each of the RGB layers. SdRGB was thus conceived as a useful feature for separating between these two general colour groups
RedCHROMATIC GreenCHROMATIC BlueCHROMATIC	Chromatic values for each RGB layer. Created by dividing relevant normalised band value (e.g., R for RedCHROMATIC) by the sum of all normalised RGB values. Reduces variance in pixel values due to illumination variance in the image, required for calculation of additional vegetation indices [35]
GRVI	Index for discriminating vegetative surfaces from non-vegetative background [36]. Green Red Vegetation Index = (GreenCHROMATIC − RedCHROMATIC)/(GreenCHROMATIC + RedCHROMATIC)
ExcessGREEN	Alternative vegetation index to GRVI, and provides measure of pixel green colour strength. ExcessGREEN = (2 × GreenCHROMATIC) − RedCHROMATIC − BlueCHROMATIC [35]
ExcessRED	Additional vegetation index that measures excess Red in pixel. ExcessRED = (1.4 × RedCHROMATIC) − GreenCHROMATIC [35]
ExGREENminusExRED	Alternative vegetation index [32]. ExGREENminusExRED = ExcessGREEN − ExcessRED
PCA1 PCA2	3 × Principal Components (PCA) features created to reveal hidden variance in relationships between RGB layers. First two layers (PCA1 and PCA2) found to contain useful information and retained for further analysis
PCA_DIFF	Investigation indicated some potential differences between Smooth vegetation (typically grass) and rough surface vegetation (trees and shrubs) surfaces in layer values for both PCA1 and PCA2. PCA_DIFF = |PCA1 − PCA2|
Object—Based Feature	*Description*

Table 3. *Cont.*

Feature Layer	Description
RVindex	PCA_DIFF pixel values tend to vary more significantly within rough surface vegetation objects than smooth vegetation objects, resulting in significant layer texture. Typical texture features within eCognition are computationally expensive to implement, therefore standard deviation of within object pixel values provides a rough approximation of texture for any given layer. Standard deviation of object PCA_DIFF values are further normalised with object ExcessGREEN values, as rough surface vegetation objects have higher values in this feature than smooth vegetation objects. $RVindex_{Object}$ = $Mean.ExGreen_{Object} \cdot Standard.Deviation(PCAdiff)_{Object}$
REDminusBLUE	Simple arithmetic feature to provide estimation of object browness. $REDminusBLUE_{Object}$ = $Mean.Red_{Object} - Mean.Blue_{Object}$
Brightness	Default software feature calculated from image layers (see [37] p.233)

Separation into Super-Classes

Multi-resolution segmentation of the MeanRGB layer (input parameters: Scale Factor = 1; Shape = 0.1; Compactness = 0.1) created the initial image segments. The parameters used result in over-segmentation, and were designed to ensure a minimum of interclass object overlap within small surface patches found in garden areas. This in turn provided spectrally pure seed objects to initiate the topological processing algorithms described below. The resulting objects were classified according to membership of either the Vegetation, Manmade or Shadow superclass groups. Discrepancies in quality between the 2009/10 and 2015 image tiles required the adoption of different superclass separation rules to classify tiles for the different image acquisition dates (Supplementary Materials, Figure S2). The final stage of this routine was to clean the vegetation class of potential shadow pixel contamination. Extremely dark pixels artificially skew the RVindex, by increasing pixel variance in the PCA_DIFF feature, thus increasing confusion between Grass and Rough Surface Vegetation (Shrubs and Trees) objects.

Identification of Tree Canopy Objects

Due to spectral correlation with other vegetative classes, the identification of tree canopies required the use of the digitised tree canopy data to provide the absolute spatial extents of where tree canopies could exist in the TAI in relation to other vegetative areas. However, discrepancies between the digitised tree canopy data and canopy extents in the imagery existed due to the removal of tree canopies between acquisition dates of the tree audit data and TAI. Thus, an iterative tree canopy identification and region growing process, adapted from the method described by [38], was required to classify image canopy extents within the generalised tree audit polygon areas. The nested loop uses topological object processing, by first identifying initial tree seed objects with a high level of certainty. Tree regions are then grown through neighbouring object reassignment according to general intra-canopy inter-object spatial correlation in both the Brightness and Rvindex features (Supplementary Materials, Figures S3 and S4).

Classification of Grass and Shrubs

Due to feature similarity between sample classes representing Grass and Shrubs, a similar iterative seed placement and region growing routine was implemented to assign objects to either class. Seed placement for both Grass and Shrub seeds began with thresholds designed to identify such objects with a high degree of certainty. Intra-class spatial correlation in the RVindex and Brightness features guided the optimisation of topological processing algorithms, which then enabled region growing from seeds into neighbouring vegetation objects of the same class. The seeding and region growing routine was designed to gradually reduce uncertainty in classification and ensure all vegetation objects were assigned to either class (Supplementary Materials, Figures S5 and S6).

Classification Optimisation

The vegetation classification routines result in generally agreeable classification outputs. However, patterns of misclassification were evident, requiring further processing to re-assign objects to the correct classes (Supplementary Materials, Figure S7). Once homogenous object areas are optimised through this process, Bare Earth objects are then classified, and then assessed against a number of testing thresholds to ensure correct assignment. This is required, as Bare Earth is spectrally similar to other surfaces (e.g., roof tiles) within the manmade class. Further optimisation routines are then implemented to optimise borders between classes, and remove small insignificant object areas ($\leq$150 pixels) to relevant surrounding object classes, or surrounding larger class objects with greater than majority border relation (Supplementary Materials, Figure S8) [39].

Building Classification

Classified polygons were clipped to OSMT garden extents, with any area outside garden areas assigned as a non-garden area. Resulting class polygons were exported as vector shapefiles. Garden buildings identified previously were integrated with this data set to classify building areas. Associated polygons superseded the areas of all overlapping class polygons regardless of class. This step was undertaken, as it proved extremely difficult in the experimentation phase to distinguish building roofs from other manmade surfaces within spectral features [40].

Accuracy Assessment

As the classification ruleset had been altered to accommodate differences in image quality between image dates, confusion matrices were generated for the 2009/2010 and 2015 image data respectively. The minimum number of samples (n) required for overall accuracy assessment was determined through use of the Multinomial Law equation [41]:

$$n = B/4b^2 \tag{9}$$

where n is the overall number of samples; B is determined from the required confidence level (e.g., 95%) and is the chi-square critical value for 1 d.f and $\chi^2_{(1-\alpha/k)}$ (where k is the number of classes, α is 1 minus the required confidence level e.g., 95% = 1 − 0.95); b is the desired level of precision (e.g., 5% = 0.05). Minimum number of samples were calculated for (k = 7) classes in total, confidence level of 95% and required precision of 3%. In this instance B = 7.23 and b = 0.03, therefore:

$$n = 7.23/(4 \times 0.03^2) = 2008.33 = 2009 \tag{10}$$

Proportional class coverage varied between classes. Therefore, the total minimum sample number was stratified according to these proportions. To ensure a minimum acceptable number of classes, an additional 50 samples were added per class. As object sizes varied, class samples were stratified within class populations according to object size quantiles. Assessment objects were superimposed onto the image data, with class labels describing the majority class within the object area attached after manual visual validation.

Extrapolation within Shadow Class

A simple topological processing method (after [42]), reassigned shadow object area proportions according to relative border relationships to neighbouring object classes. As shadow is likely to cover neighbouring objects to some degree, this process provides a reasonable method to reduce redundant information in the classification. Excluded from this process was the building class, as this was not classified from the imagery directly. Remaining shadow objects, with no neighbouring class objects, were re-assigned surface proportions according to surface proportions obtained from gardens with minimal (<5%) shadow coverage.

3. Results

3.1. Citizen Science Survey Responses

In total 1031 responses were received for the Manchester Local Authority area. The distribution of these responses varied across the 32 wards with 11.4% of participants residing in the most represented ward (Chorlton) and less than 1% of participants from the least represented ward (Longsight) (Figure 5a). The proportional split between responses from different housing types was closely aligned with the proportions of terraced and detached housing in Manchester (36% and 5%, respectively) but semi-detached housing was over-represented (43%) at the expense of flats, maisonettes, and apartments (16%). Two thirds of responses were from households where the home was owned outright or with a mortgage/loan. The remainder of responses were from households in rental properties from private landlords, registered landlords/ housing associations or the council. A representative sample of household and garden types was obtained, with 24% of survey responses from households with less than 20% garden green space, compared to 20% survey responses from households with greater than 80% garden green space, thus it was not only households with large green gardens that completed the survey (Figure 5b).

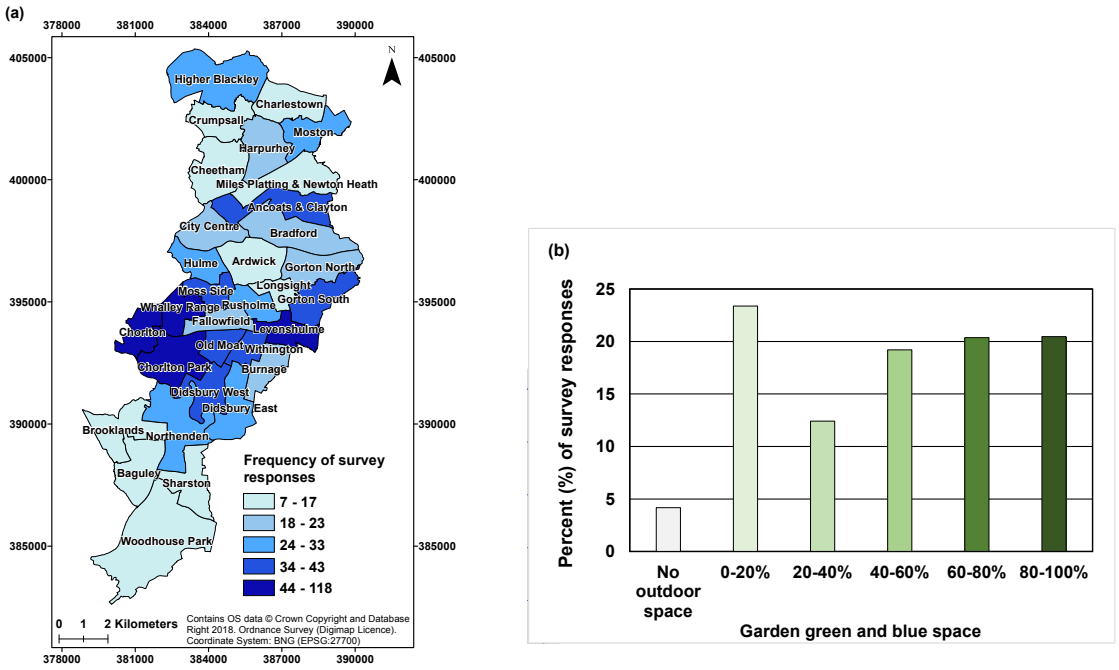


Figure 5. Citizen Science Survey responses (**a**) by ward; (**b**) by proportion of garden green and blue space.

Overall results of the garden composition derived from valid CSS response estimates (see Section 2.3.2) reveal that, on average, across Manchester, 51.88% (s = 30.62%) of a garden is comprised of green infrastructure (Table 4). There is, understandably, significant variation around these values in the CSS responses. At the extremes, a small number of responses indicated the garden was either wholly impervious (2.43%) or wholly pervious (3.45%). Of the ten surface cover categories, hard impervious and mown grass were the dominant features of gardens, occupying an average of 26.64% and 20.79% of a garden's area, respectively. In contrast, rough grass and water were the least present features, accounting for 2.95% and less than 1% (0.61%) of an individual garden area.

Table 4. Proportional surface coverage (%) of an average Manchester garden from the CSS responses with valid estimation range (n = 758). All categories reported minimum values of zero.

CSS Surface	Mean	SD	Max.
Buildings	5.85	10.74	81.21
Bare soil	8.74	14.15	100
Cultivated	11.82	14.31	77.14
Hard impervious	26.64	28.15	100
Hard pervious	6.89	15.59	100
Mown grass	20.79	22.78	94.09
Rough grass	2.95	10.66	100
Shrubs	10.94	11.46	74.35
Trees	4.77	7.34	50.91
Water	0.61	2.07	26.59
Green Infrastructure *	51.88	30.62	100

* *Green Infrastructure = Cultivated + Mown Grass + Rough Grass + Shrubs + Trees + Water.*

Average garden size from the survey responses indicated increases from 84.81 m^2 (s = 103.23 m^2) to 205.68 m^2 (s = 208.89 m^2) to 427.82 m^2 (s = 789.69 m^2) for terraced, semi-detached and detached housing, respectively. These estimations are higher than those found in Sheffield, UK [43], reflecting the shift in focus from the rear garden to the total garden area. The greatest discrepancy between the values quoted here and those in Sheffield, UK [43] is for detached housing, where housing frontages tend to be larger.

3.2. Validation of Garden Composition Derived from the Citizen Science Survey

Out of the total 1031 survey responses, 758 (73.5% of the total) were identified as having valid surface estimation ranges (see Section 2.3.2), with a further 728 correctly identified with relevant OSG polygon area. Mean VA was calculated as 76.63% (95% confidence interval: 78.53%/74.73%), s = 15.24%, with Median VA of 79.65%. This is encouraging given that 85% accuracy is general considered as the benchmark for classification exercises [31]. In addition, it is likely that limitations associated with the validation process impacted upon the VA. This was primarily due to potential temporal differences of up to 7 years between the survey date (2016) and TAI (2009–2015), and therefore, any changes to garden composition within this timeframe may have resulted in under-reporting of VA. In addition, shadow and angular obscuration issues within the imagery may have introduced error in the VA estimation due to the subjective digitisation and categorisation of TAI.

As limited guidance was given to survey respondents on how to estimate their garden surface composition, two sets of correlation analysis were undertaken to assess if VA was associated with any particular land surface types from the digitised validation exercise or CSS estimation. The first set was conducted on all validation records. The second set controlled for missing values in each variable i.e., where a certain validation surface group was not included in CSS surface estimation (Table 5).

Table 5. Association between VA and land surface types estimated from digitisation and CSS (ns = $p \geq 0.05$; * = $p < 0.05$; ** = $p < 0.01$; *** = $p < 0.001$; **** = $p < 0.0001$).

Land Surface Type Correlated to VA	Set 1		Set 2	
	Cor	**(*p-Value*)**	**Cor**	**(*p-Value*)**
Proportion of digitised land surface area as VSG				
DIG.Buildings	0.13	*ns*	0.07	*ns*
DIG.Grass	0.13	*ns*	−0.1	*ns*
DIG.Manmade	0.22	***	0.22	***
DIG.Shrubs	0.17	**	0.13	*
DIG.Trees	−0.48	****	−0.48	****
Proportion of CSS estimated land surface area as VSG				
VSG.Buildings	0.08	*ns*	0	*ns*
VSG.Grass	−0.26	***	−0.34	****
VSG.Manmade	0.28	****	0.29	****
VSG.Shrubs	0.07	*ns*	0.01	*ns*
VSG.Trees	0.07	*ns*	−0.06	*ns*
Total digitised garden area	−0.21	***	*N.A*	- - -

Analysis reveals generally weak correlation values for statistically significant analyses. VSG variables assessed whether CSS respondents may have had difficulty in estimating certain surface types. DIG.Trees exhibited a moderate negative association with VA for both sets of analysis. It is possible that respondents considered tree coverage only for within garden trees (i.e., with tree stand in actual garden limits), thus ignoring external overlapping tree coverage. This result also suggests some ambiguity in respondent interpretation of tree surface coverage estimation, according to whether surface coverage is represented by canopy or tree stand extents. However, as the variables are proportional and thus dependent on each other, it is difficult to isolate the exact impact on misclassification of VSG estimation from one category from to another. It was theorised that possible difficulties in respondent CSS may have arisen with increasing garden areal extents (total digitised garden area), however, *cor* = −0.21, and therefore, this is not supported in the above analysis.

Image Classification

The overall accuracy differed slightly between the assessments for both image dates. The classifications are close to the general classification acceptability, with low confidence estimates falling just below the 85% accuracy threshold; 83.22% and 81.92% for 2009/2010 and 2015 assessments respectively (Tables 6 and 7). Kappa values for both assessments indicate a strong likelihood that assessment agreement is not due to chance [41]. Figure 6 provides a sample of the image classification output.

Table 6. Confusion Matrix for 2009/2010 image classification.

Land Surface Type	Bare Earth	Buildings	Grass	Manmade	Shadow	Shrubs	Trees	User's Accuracy
Bare Earth	**28**	0	9	6	0	10	3	50.0
Buildings	0	**65**	2	1	1	3	4	85.5
Grass	2	1	**304**	4	0	17	4	91.6
Manmade	12	110	4	**503**	0	10	4	78.2
Shadow	0	5	0	3	**576**	4	4	97.3
Shrubs	1	0	57	5	2	**291**	16	78.2
Trees	0	0	4	0	1	16	**267**	92.7
Producer's Accuracy	65.1	35.9	80.0	96.4	99.3	82.9	88.4	**2359**
Overall Accuracy:	86.22% (High 89.22%/**Low 83.22%**)							
Kappa:	0.831							

Table 7. Confusion Matrix for 2015 image classification.

Land Surface Type	Bare Earth	Buildings	Grass	Manmade	Shadow	Shrubs	Trees	User's Accuracy
Bare Earth	**50**	0	2	15	0	0	1	73.5
Buildings	0	**67**	1	4	0	1	3	88.2
Grass	7	0	**270**	1	4	43	15	79.4
Manmade	16	67	4	**623**	13	8	5	84.6
Shadow	0	0	0	4	**491**	2	3	98.2
Shrubs	3	1	24	7	38	**245**	18	72.9
Trees	1	0	7	1	14	23	**258**	84.9
Producer's Accuracy	64.9	49.6	87.7	95.1	87.7	76.1	85.1	**2360**
Overall Accuracy:	84.92% (High 87.92%/**Low 81.92%**)							
Kappa:	0.813							

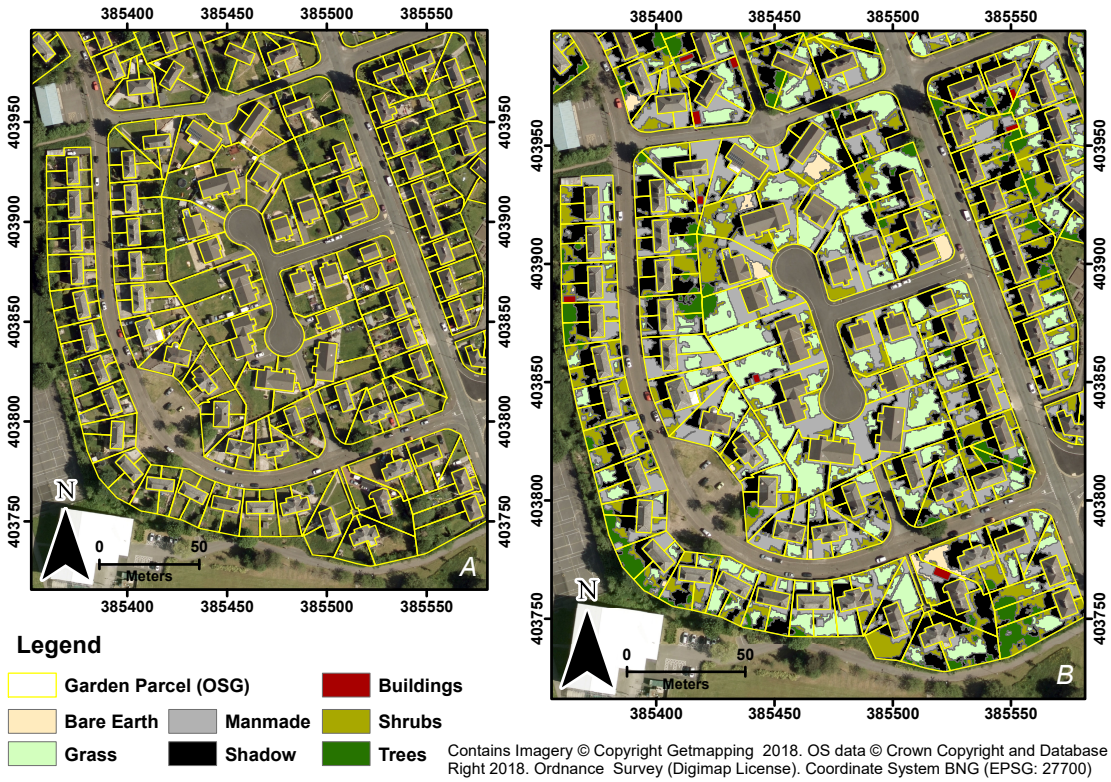


Figure 6. Example results of the image classification (**A**) original image; (**B**) classification output.

Interclass accuracy varies significantly between classes. User's accuracy (UA) determines the level of map accuracy for the end user, by estimating the accuracy that a given class area in the map represents the relevant class within the image data [41]. Accuracy estimation should return relative estimation parity between classes in order to provide a useful approximation of overall estimation [44]. However, class accuracy estimations deviated from this ideal for a number of reasons outlined below.

Bare Earth performed poorly in the 2009/2010 image assessment with 50% UA, with confusion of 17.86%, 16.07%, 10.71% and 5.36% with the Shrubs, Grass, Manmade and Trees classes respectively. In the 2015 image assessment, Bare Earth performed better with 73.5% UA, and with 22.06%, 2.94% and 1.47% confusion with the Manmade, Grass and Trees classes respectively. Bare Earth is spectrally similar to other manmade class areas (i.e., clay roof, brown decking areas), and this resulted in inaccurate feature thresholds for this class, particularly for the 2009/2010 data.

UA between image dates was less variable for other land cover classes. UA for Shrubs was 78.2% (2009/2010) and 72.9% (2015), with intra-vegetation superclass confusion apparent for both

image data sets. 15.32% (2009/2010) and 7.14% (2015) of classified Shrubs objects were identified as Grass in assessment, with a further 4.3% (2009/2010) and 5.36% (2015) identified as Trees. Grass with UA of 91.6% (2009/2010) and 79.4% (2015) exhibited similar patterns of vegetation classification confusion. The use of ancillary data aided the region growing process for the Trees class, by limiting the domain of region growing within definitive confines. With UA values of 92.7% (2009/2010) and 84.9% (2015), Trees performed better than the other vegetative classes. The vegetation indices derived from the literature proved useful in discriminating between the vegetation classes [35,36] in the appropriate routines.

The Manmade class achieved UA of 78.2% (2009/2010) and 84.6% (2015). The major source of class confusion, for both sets of image data, was with the Buildings class, which accounted for 17.11% (2009/2010) and 9.1% (2015) of manmade misclassification. The TAI imagery is collected from slightly off-nadir angles, with vertical facades/rooftops causing some obscuration of associated ground features in the imagery. OSG data, representing garden parcel extents on the ground level, was found to overlap image features, resulting in unwanted inclusion of building features within garden parcels. This is indicated by low producer's accuracy for the Building class of 35.9% (2009/2010) and 49.6% (2015).

The Buildings class achieved UA of 85.5% (2009/2010) and 88.2% (2015). Misclassifications here again result largely due to inconsistencies between TAI and OSG data. Garden Buildings may be present but are subsequently hidden under covering tree canopies in the imagery, thus preventing confirmation of whether building features are present. In addition, due to temporal differences between the OSG and TAI, building objects may be represented in OSMT data that were not yet constructed by the time of TAI image acquisition.

The effect of extrapolation within the shadow class upon the classified area proportions is shown in Figure 7. In the Original Classification (OC), 23.27% of the total classified area is assigned as *Shadow*. The extrapolation method used in this study does not ascribe class areas spatially within *Shadow* proportions, therefore it proved difficult to verify the results of this study with the TAI. Methods similar to the CSS validation could have been employed, however, this would bias assessments towards extrapolated proportions (see Section 2.3.3). Class proportions calculated to total classified areas after SL2 processing (TP), were compared to class proportions extrapolated from those calculated within the non-Shadow area (SA) representing simple weighted-up proportions from OC. In both extrapolation methods, building proportions remained absolute (see Section 2.3.3). Average class proportion difference between TP and SA was 0.96%, indicating the limited effect of extrapolation on results obtained solely from OC proportions. Shrubs and Trees TP proportions are higher than SA, as these object areas are likely to cast shadow in the imagery and thus neighbour shadow proportionally than ground surfaces.

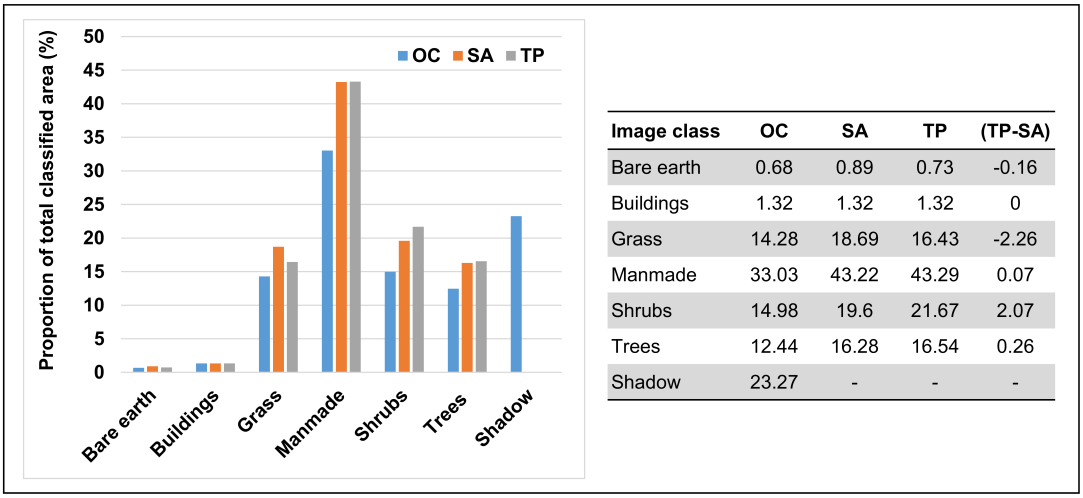


Image class	OC	SA	TP	(TP-SA)
Bare earth	0.68	0.89	0.73	-0.16
Buildings	1.32	1.32	1.32	0
Grass	14.28	18.69	16.43	-2.26
Manmade	33.03	43.22	43.29	0.07
Shrubs	14.98	19.6	21.67	2.07
Trees	12.44	16.28	16.54	0.26
Shadow	23.27	-	-	-

Figure 7. Comparison of total classification class proportions (%) between OC (*blue*), SA (*orange*) and TP (*grey*) methods of class proportion coverage calculation.

3.3. Extrapolation of CSS Surface Proportions to Classified Data

Extrapolation of CSS surface class proportions within SL2 classified garden areas, required an estimated minimum of 383 samples per surface class (95% confidence level, assuming maximum variance of surface estimations for each surface sample) to the total number of 156,573 OSG garden parcels. However, due to the difficulty in accounting for issues associated with self-selection bias in the CSS survey, extrapolation was-based upon the entire population of responses with valid surface estimation ranges (n = 758).

Extrapolation multipliers were therefore calculated by first finding the image class areas of total garden area obtained from the CSS responses, and then calculating CSS surface proportions to relevant class area (Table 8). Variation between the total extrapolated surface coverage proportions of each CSS surface for all OSG areas, in comparison to CSS estimates is shown in Table 9. Variation between estimated CSS surfaces, and extrapolated surface proportions is evident for all classes. CSS estimates are lower for the Hard Impervious, Hard Pervious and Trees class, while the reverse is true for all other classes. Interestingly, the estimated OSG green infrastructure coverage is less than the overall CSS estimates, with a 10% difference between the CSS proportions to CSS garden area, and CSS proportions to total garden area.

Table 8. Extrapolation multipliers for CSS surfaces calculated as within image classification class proportion.

	Image Classification Class							
	Grass		**Manmade**		**Shrubs and Bare Earth**			
CSS Surfaces to Image Classification class	Mown Grass	Rough Grass	Hard Impervious	Hard Pervious	Bare Earth	Cultivated	Shrubs	Water
Within class proportion	0.88	0.12	0.78	0.22	0.23	0.34	0.41	0.02

Table 9. Comparison between CSS surface proportion estimates, CSS surface proportions per total CSS garden area and extrapolated CSS proportions within final classification.

Land Surface Cover Type	**Mean Reported CSS Proportions for ALL Responses**	**CSS Surface Proportions per Total CSS Garden Area**	**CSS Proportions to Total OSG Garden Area**
Bare Soil	8.74	7.63	5.15
Buildings	5.85	5.58	1.32
Cultivated	11.82	11.28	7.62
Hard Impervious	26.64	19.65	33.77
Hard Pervious	6.89	5.54	9.53
Mown Grass	20.79	24.97	14.46
Rough Grass	2.95	3.40	1.97
Shrubs	10.94	13.6	9.19
Trees	4.77	7.69	16.54
Water	0.61	0.66	0.45
Green infrastructure *	51.88	61.60	50.23

* *Green infrastructure = Cultivated + Mown Grass + Rough Grass + Shrubs + Trees + Water.*

3.4. Green Infrastructure (GI) in Manchester's Gardens

Overall, Manchester has 23.63 km^2 of gardens which make up 20.4% of Manchester's land area, yet just 11.87 km^2 of this land area is GI. There is also substantial spatial variation evident in GI across Manchester, with garden GI ranging between 0–27% across Manchester's wards (Figure 8). While wards with the greatest proportion of garden space (including Old Moat, Withington and Burnage) (Figure 1) still have the highest garden GI as a percentage of ward area, they experience a significant drop in garden GI, of almost half the amount (for example, Burnage domestic gardens cover 47.2% of the ward area, but garden GI coverage is only 27% of the ward area). This demonstrates that gardens are not performing as well as they could be in delivering GI. Furthermore, comparing this to public green infrastructure data [45] indicates that Manchester has, in total, 49.0% GI (equating to 56.66 km^2), and 20.94% of this GI is contained within domestic gardens. City residents, therefore, are responsible for the

maintenance and safeguarding of around one fifth of the city's GI. Furthermore, private gardens can contribute from 1% up to 62% of the overall GI within a ward (Figure 9).

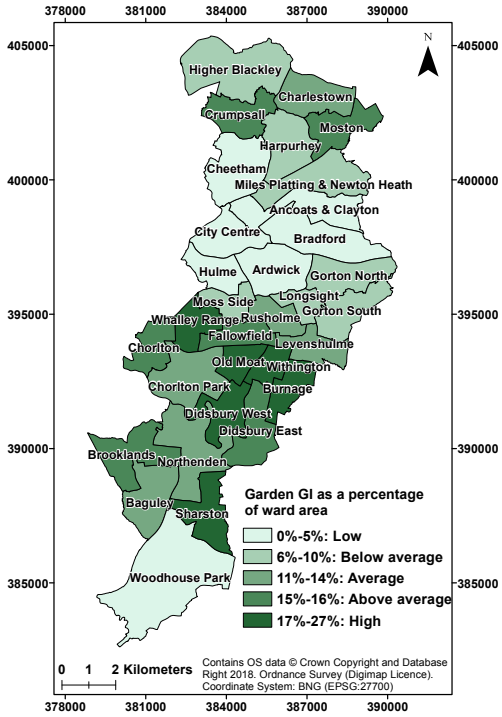


Figure 8. Garden green infrastructure (GI) as a percentage of ward area (classified by quantiles).

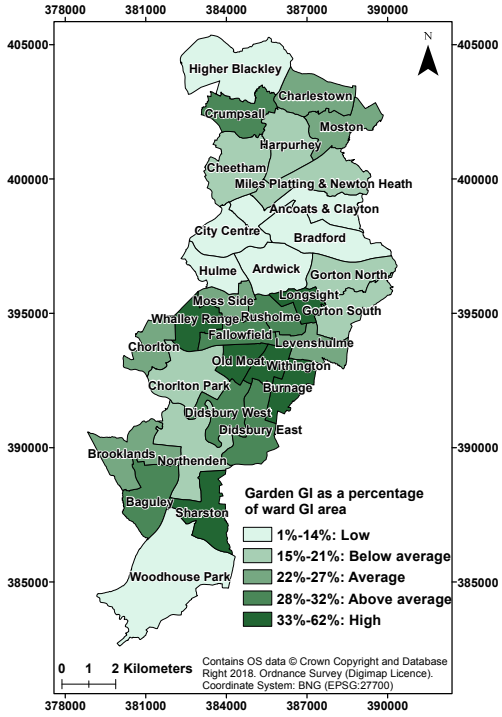


Figure 9. Garden green infrastructure (GI) as a percentage of ward GI area (classified by quantiles).

4. Discussion

A representative sample of household and garden types was gained from the 1031 respondents to the citizen science survey in Manchester. The number of responses to the survey, however, varied greatly across wards, with wards in north and south Manchester having less than twenty responses, highlighting the need for a robust methodology when using citizen science data to validate and extend the database beyond areas with high response rates.

Validation of the garden land surface composition estimated by CSS respondents revealed an average validation accuracy of 76.63% (s = 15.24%), close to the benchmark of 85% accuracy for classification exercises [31]. This figure indicates that citizens are able to reasonably estimate the proportional land surface coverage within their gardens. Differences between the garden composition obtained from the survey responses and classified aerial imagery may have been caused by a number of issues. Firstly, there was a temporal difference of up to 7 years between the survey date (2016) and TAI (2009–2015). Changes to garden composition within this timeframe may have introduced errors in the VA. This also highlights the benefit of obtaining information from citizens—it is more up-to-date than other data sources, and may therefore be more accurate. Secondly, limited guidance was provided to survey respondents about exactly how to estimate garden composition. Estimating tree coverage may be particularly challenging to undertake without guidance, since it could be considered as the tree canopy or the surface cover of a tree at ground level. Tree surfaces were considered as tree canopy in the digitisation process, given that the ultimate aim was to investigate ecosystem service benefits, and trees are more productive than other surface categories. Thirdly, errors may have been introduced by the subjective digitisation and categorisation of the TAI [46].

Validation of CSS responses indicated that they yielded useable information for further extrapolation purposes. While the validation methods adopted here are not without error, conducting controlled field surveys to map exact garden surface coverages on this large scale would be a significant undertaking, with researchers required to engage with multiple garden owners for access [1]. Imagery interpretation thus provides a cost and time effective method for validating such information [23,47]. As the CSS was designed to be simple and accessible for the general population, limited information was provided on how to assess garden surfaces. Inclusion of additional material for respondents may have improved general accuracy of responses. However, increasing the burden of the task may have discouraged potential respondents. Therefore, the trade-off between accuracy and accessibility in designing the CSS, was successful in obtaining a good sample of useable garden surface estimations.

The image classification produced reasonable results, with variation in garden surfaces across Manchester mapped to a high level of detail. A significant proportion of the original classification (23.27%), however, required extrapolation within obscured shadow areas. Overall accuracy of final image surface estimation is therefore expected to be lower than overall accuracies reported, as additional error is introduced through the shadow extrapolation method, which is based on assumptions of a shadow areas' relationship to its neighbouring class surfaces. Shadow compensation techniques exist for high resolution imagery, but are computationally expensive to implement, and may only provide marginal gains in information content for classification purposes [48,49]. The spatial extrapolation method is quick to implement during post-classification processing, however the lack of validation results in unquantified uncertainty in the final extrapolation results.

The classification methods used in this study were constrained by the computational limitations associated with processing the available image data. The classification routines in this study were designed and optimized based on experimentation. Machine learning classification techniques such as Support Vector Machines, Random Forests enable quick and effective classification of high resolution imagery [50–52], however, the limited spectral resolution of the aerial imagery resulted in considerable feature overlap between image classes, thus inhibiting the use of such methods. The region growing and cleaning routines compensated for this, by instead using spatial autocorrelation between image objects, with neighbor to neighbor spatial and spectral similarity features accounting for fuzziness between image classes [34,53]. A limitation of this study is that the classification routines were

developed from personal interpretation by the image analyst, with such knowledge not easily transferrable to other exercises. Formalization of the ontology of object types and relationships used in classification schemes therefore may improve transferability of such knowledge within the OBIA classification domain [54].

Extrapolation between CSS estimates and the classified imagery was conducted using a simple method. In other urban studies, surveyed garden/parcel vegetation characteristics, obtained from a limited sample, are used to infer findings to wider urban areas according to parcel landcover-landownership [55], parcel land-value [56] and housing type [57] categories. As the level of garden vegetative cover within gardens has been found to be statistically associated with such variables, then stratifying the CSS valid surface estimation records into such categories may arguably improve the extrapolation process. However, as the CSS estimates are likely to contain a certain degree of error (due to user assessment and self-selection bias), and as the validation exercise highlighted no significant variables to predict CSS error, based on general statistical assumptions the entire valid CSS sample was used to include as much CSS information in the extrapolation process as possible. The impact of sources of uncertainty in both datasets is, therefore, not explicitly accounted for in the final extrapolated CSS surface garden outputs. The overall approach could be further improved to incorporate sensitivity and uncertainty analysis, to indicate both sources of error in the methodology, and provide useable estimates of error for end users of the data [58].

Overall, the large differences between the estimated CSS surface proportions and final extrapolated OSG surface proportions for the whole study area indicate the limitations of simple extrapolation from a limited sample of survey responses (Table 9), supporting the need for an extrapolation approach as developed in this study. Heterogeneity in garden surface composition is difficult to estimate from a biased survey sample. As evident from the results, CSS survey estimates of green infrastructure cover differ by up to 10% (dependent on calculation method) compared to final OSG extrapolated figures. A simple unweighted extrapolation from the survey data therefore may overestimate city-wide garden green infrastructure coverage considerably. The benefits of combining the two datasets is evident by this fact, as spatial heterogeneity of garden surfaces is measured from the classified data.

5. Conclusions

A combined approach was applied to classify land surface cover of urban domestic gardens using citizen science data and high resolution image analysis. This combined approach was advantageous for two main reasons. Firstly, it enabled an assessment of the quality of citizen science data collected in relation to estimations of land cover proportions of ten common domestic garden land surface cover types, finding that citizens are able to quite accurately estimate this within their gardens (mean Validation Accuracy = 76.63%, s = 15.24%). Secondly, this approach enabled extending an object-based image classification to extrapolate the ten land surface cover types within gardens across the whole study area. Furthermore, engagement of the public in the citizen science survey provides more up-to-date data than aerial imagery and facilitates education for sustainable development, specifically through informing the public about the value of urban gardens in the provision of urban green infrastructure and its associated ecosystem services.

The final dataset reveals that domestic gardens contain 20.94% of Manchester's GI, and there is clear spatial variation across the study area. Such privately owned land is perceived by policymakers and urban planners as challenging to influence. Detailed evidence on the current urban GI in gardens is valuable for urban planning stakeholders in the local area in order to deliver targeted GI interventions within and beyond gardens, for example, through identification of areas of GI need within neighbourhoods. The detailed land surface classification information is useful for a broad range of analyses. For example, distinguishing between pervious and non-pervious surfaces within the manmade class enables better estimation of the flood attenuation function of local garden areas from a simple vegetative/non-vegetative classification [2,8].

Supplementary Materials: The following are available online at http://www.mdpi.com/2072-4292/10/4/537/s1. Figure S1. Extent of the True-colour Aerial Imagery (TAI) for Manchester, UK; Figure S2. Thresholds separate initial image segments into superclass groups; Figure S3. Tree classification routine overview; Figure S4. Tree classification region growing; Figure S5. Grass and Shrubs classification—seed placement; Figure S6. Grass and shrubs classification Figure S7. Rules for classification class cleaning; Figure S8. Classification routines to optimise borders between classes.

Acknowledgments: This work was carried out as part of the Assessing the contribution of domestic gardens to urban ecosystem services project (2016–2018), funded by Natural Environment Research Council NE/N017374/1 and NE/N017374/2. We gratefully acknowledge support from the project team and project partners.

Author Contributions: Fraser Baker, Claire Smith, and Gina Cavan conceived and designed the work and wrote the manuscript. Fraser Baker performed the data processing and image classification. Fraser Baker and Gina Cavan performed the GIS analysis. Gina Cavan provided final editing on the paper.

Conflicts of Interest: The authors declare no conflict of interest. The founding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, and in the decision to publish the results.

References

1. Mathieu, R.; Freeman, C.; Aryal, J. Mapping private gardens in urban areas using object-oriented techniques and very high resolution imagery. *Landsc. Urban Plan.* **2007**, *81*, 179–192. [CrossRef]
2. Warhurst, J.R.; Parks, K.E.; McCulloch, L.; Hudson, M.D. Front gardens to car parks: Changes in garden permeability and effects on flood regulation. *Sci. Total Environ.* **2014**, *485*, 329–339. [CrossRef] [PubMed]
3. Cameron, R.W.F.; Blanuša, T.; Taylor, J.E.; Salisbury, A.; Halstead, A.J.; Henricot, B.; Thompson, K. The domestic garden–Its contribution to urban green infrastructure. *Urban For. Urban Green.* **2012**, *11*, 129–137. [CrossRef]
4. Newman, G.; Chandler, M.; Clyde, M.; McGreavy, B.; Haklay, M.; Ballard, H.; Gray, S.; Scarpino, R.; Hauptfeld, R.; Mellor, D.; et al. Leveraging the power of place in citizen science for effective conservation decision making. *Biol. Conserv.* **2017**, *208*, 55–64. [CrossRef]
5. Loram, A.; Warren, P.; Thompson, K.; Gaston, K. Urban domestic gardens: The effects of human interventions on garden composition. *Environ. Manag.* **2011**, *48*, 808–824. [CrossRef] [PubMed]
6. Lin, B.B.; Gaston, K.J.; Fuller, R.A.; Wu, D.; Bush, R.; Shanahan, D.F. How green is your garden? Urban form and socio-demographic factors influence yard vegetation, visitation, and ecosystem service benefits. *Landsc. Urban Plan.* **2017**, *157*, 239–246. [CrossRef]
7. DCLG, Department for Communities and Local Government (DCLG). Guidance on the Permeable Surfacing of Front Gardens. 2008. Available online: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/7728/pavingfrontgardens.pdf (accessed on 12 February 2018).
8. Perry, T.; Nawaz, R. An investigation into the extents and impacts of hard surfacing of domestic gardens in an area of Leeds, United Kingdom. *Landsc. Urban Plan.* **2008**, *86*, 1–13. [CrossRef]
9. Ghosh, S.; Head, L. Retrofitting the suburban garden: Morphologies and some elements of two Australian residential suburbs compared. *Aust. Geogr.* **2009**, *40*, 319–346. [CrossRef]
10. Sayce, S.; Walford, N.; Garside, P. Residential development on gardens in England: Their role in providing sustainable housing supply. *Land Use Policy* **2012**, *29*, 771–780. [CrossRef]
11. Hof, A.; Wolf, N. Estimating potential outdoor water consumption in private and urban landscapes by coupling high-resolution image analysis, irrigation water needs and evaporation estimation in Spain. *Landsc. Urban Plan.* **2014**, *123*, 61–72. [CrossRef]
12. Cameron, R.W.; Blanusa, T. Green infrastructure and ecosystem services—Is the devil in the detail? *Ann. Bot. Lond.* **2016**, *118*, 377–391. [CrossRef] [PubMed]
13. Gill, S.E.; Handley, J.F.; Ennos, A.R.; Pauleit, S. Adapting cities for climate change: The role of green infrastructure. *Built Environ.* **2007**, *33*, 115–133. [CrossRef]
14. Hall, J.M.; Handley, J.F.; Ennos, A.R. The potential of tree planting to climate proof high density residential areas in Manchester, UK. *Landsc. Urban Plan.* **2012**, *104*, 410–417. [CrossRef]
15. Armson, D.; Stringer, P.; Ennos, A.R. The effect of tree shade and grass on surface and globe temperatures in an urban area. *Urban For. Urban Green.* **2012**, *11*, 245–255. [CrossRef]

16. Dennis, M.; Barlow, D.; Cavan, G.; Cook, P.A.; Gilchrist, A.; Handley, J.; James, P.; Thompson, J.; Tzoulas, K.; Wheater, C.P.; et al. Mapping urban green infrastructure: A novel landscape-based approach to incorporating land use and land cover in the mapping of human-dominated systems. *Land* **2018**, *7*, 17. [CrossRef]
17. Kazmierczak, A.; Cavan, G. Surface water flooding risk to urban communities: Analysis of vulnerability, hazard & exposure. *Landsc. Urban Plan.* **2011**, *103*, 185–197.
18. Smith, C.L.; Lawson, N. Identifying Extreme Event Climate Thresholds for Greater Manchester, UK: Examining the Past to Prepare for the Future. *Meteorol. Appl.* **2012**, *19*, 26–35. [CrossRef]
19. Carter, J.G.; Cavan, G.; Connelly, A.; Guy, S.; Handley, J.; Kazmierczak, A. Climate change and the city: Building capacity for urban adaptation. *Prog. Plan.* **2014**, *95*, 1–66. [CrossRef]
20. Levermore, G.J.; Parkinson, J.B.; Laycock, P.J.; Lindley, S.J. The Urban Heat Island in Manchester 1996–2011. *Build. Serv. Eng. Res. Technol.* **2015**, *36*, 343–356. [CrossRef]
21. Intelligence Hub (Manchester Statistics) Mid-Year Population Estimates for Manchester, ONS, Crown Copyright. Available online: https://dashboards.instantatlas.com/viewer/report?appid=9853f1a1b412404b806dab4d1918def7 (accessed on 12 February 2018).
22. AGMA, Association of Greater Manchester Authorities. Greater Manchester Strategic Housing Market Assessment, 2010. Available online: http://www.manchester.gov.uk/download/downloads/id/14074/gm_strategic_housing_market_assessment_shma_update_may_2010.pdf (accessed on 12 February 2018).
23. Brown, G. An empirical evaluation of the spatial accuracy of public participation GIS (PPGIS) data. *Appl. Geogr.* **2012**, *34*, 289–294. [CrossRef]
24. Getmapping, Aerial Data—High Resolution Imagery. Available online: http://www.getmapping.com/products-and-services/aerial-imagery-data/aerial-data-high-resolution-imagery (accessed on 29 March 2018).
25. City of Trees; Greater Manchester Tree Audit, Manchester, UK. Personal communication, 2011.
26. Novack, T.; Esch, T.; Kux, H.; Stilla, U. Machine learning comparison between WorldView-2 and QuickBird-2-simulated imagery regarding object-based urban land cover classification. *Remote Sens.* **2011**, *3*, 2263–2282. [CrossRef]
27. Lu, D.; Weng, Q. A survey of image classification methods and techniques for improving classification performance. *Int. J. Remote Sens.* **2007**, *28*, 823–870. [CrossRef]
28. Salas, E.A.L.; Boykin, K.G.; Valdez, R. Multispectral and texture feature application in image-object analysis of summer vegetation in Eastern Tajikistan Pamirs. *Remote Sens.* **2016**, *8*, 78. [CrossRef]
29. Blaschke, T.; Hay, G.J.; Kelly, M.; Lang, S.; Hofmann, P.; Addink, E.; Feitosa, R.Q.; van der Meer, F.; van der Werff, G.; van Coillie, F.; et al. Geographic object-based image analysis–towards a new paradigm. *ISPRS J. Photogramm.* **2014**, *87*, 180–191. [CrossRef] [PubMed]
30. Aksoy, B.; Ercanoglu, M. Landslide identification and classification by object-based image analysis and fuzzy logic: An example from the Azdavay region (Kastamonu, Turkey). *Comput. Geosci.* **2012**, *38*, 87–98. [CrossRef]
31. Foody, G.M. Harshness in image classification accuracy assessment. *Int. J. Remote Sens.* **2008**, *29*, 3137–3158. [CrossRef]
32. Huang, X.; Zhang, L. An SVM ensemble approach combining spectral, structural, and semantic features for the classification of high-resolution remotely sensed imagery. *IEEE Trans. Geosci. Remote* **2013**, *51*, 257–272. [CrossRef]
33. Feng, Q.; Liu, J.; Gong, J. UAV remote sensing for urban vegetation mapping using random forest and texture analysis. *Remote Sens.* **2015**, *7*, 1074–1094. [CrossRef]
34. Shackelford, A.K.; Davis, C.H. Fully automated road network extraction from high-resolution satellite multispectral imagery. In Proceedings of the IGARSS'03, 2003 IEEE International Geoscience and Remote Sensing Symposium, Toulouse, France, 21–25 July 2003; IEEE: Piscataway, NJ, USA, 2003; Volume 1, pp. 461–463.
35. Meyer, G.E.; Neto, J.C. Verification of color vegetation indices for automated crop imaging applications. *Comput. Electron. Agr.* **2008**, *63*, 282–293. [CrossRef]
36. Motohka, T.; Nasahara, K.N.; Oguma, H.; Tsuchida, S. Applicability of green-red vegetation index for remote sensing of vegetation phenology. *Remote Sens.* **2010**, *2*, 2369–2387. [CrossRef]
37. Definiens. Definiens Developer User Guide XD 2.0.4. 2012. Available online: www.imperial.ac.uk/media/imperial.../Definiens-Developer-User-Guide-XD-2.0.4.pdf (accessed on 18 February 2018).

38. Bunting, P.; Lucas, R. The delineation of tree crowns in Australian mixed species forests using hyperspectral Compact Airborne Spectrographic Imager (CASI) data. *Remote Sens. Environ.* **2006**, *101*, 230–248. [CrossRef]
39. Rampi, L.P.; Knight, J.F.; Pelletier, K.C. Wetland mapping in the upper midwest United States. *Photogramm. Eng. Remote Sens.* **2014**, *80*, 439–448. [CrossRef]
40. Zhou, W.; Troy, A. An object-oriented approach for analysing and characterizing urban landscape at the parcel level. *Int. J. Remote Sens.* **2008**, *29*, 3119–3135.
41. Congalton, R.G.; Green, K. *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices*; CRC Press: Boca Raton, FL, USA, 2008.
42. Zhou, W.; Huang, G.; Troy, A.; Cadenasso, M.L. Object-based land cover classification of shaded areas in high spatial resolution imagery of urban areas: A comparison study. *Remote Sens. Environ.* **2009**, *113*, 1769–1777. [CrossRef]
43. Smith, R.M.; Gaston, K.J.; Warren, P.H.; Thompson, K. Urban domestic gardens (V): Relationships between landcover composition, housing and landscape. *Landsc. Ecol.* **2005**, *20*, 235–253. [CrossRef]
44. Anderson, J.R. *A Land Use and Land Cover Classification System for Use with Remote Sensor Data*; US Government Printing Office: Washington, DC, USA, 1976; Volume 964.
45. Manchester City Council; Green Infrastructure Dataset, Manchester, UK. Personal communication, 2016.
46. Stehman, S.V.; Wickham, J.D. Pixels, blocks of pixels, and polygons: Choosing a spatial unit for thematic accuracy assessment. *Remote Sens. Environ.* **2011**, *115*, 3044–3055. [CrossRef]
47. Comber, A.; See, L.; Fritz, S.; Van der Velde, M.; Perger, C.; Foody, G. Using control data to determine the reliability of volunteered geographic information about land cover. *Int. J. Appl. Earth Obs.* **2013**, *23*, 37–48. [CrossRef]
48. Shahtahmassebi, A.; Yang, N.; Wang, K.; Moore, N.; Shen, Z. Review of shadow detection and de-shadowing methods in remote sensing. *Chin. Geogr. Sci.* **2013**, *23*, 403–420. [CrossRef]
49. Yang, J.; He, Y.; Caspersen, J. Fully constrained linear spectral unmixing based global shadow compensation for high resolution satellite imagery of urban areas. *Int. J. Appl. Earth Obs. Geoinform.* **2015**, *38*, 88–98. [CrossRef]
50. Mountrakis, G.; Im, J.; Ogole, C. Support vector machines in remote sensing: A review. *ISPRS J. Photogramm.* **2011**, *66*, 247–259. [CrossRef]
51. Duro, D.C.; Franklin, S.E.; Dubé, M.G. A comparison of pixel-based and object-based image analysis with selected machine learning algorithms for the classification of agricultural landscapes using SPOT-5 HRG imagery. *Remote Sens. Environ.* **2012**, *118*, 259–272. [CrossRef]
52. Belgiu, M.; Drăguţ, L. Random forest in remote sensing: A review of applications and future directions. *ISPRS J. Photogramm.* **2016**, *114*, 24–31. [CrossRef]
53. Burnett, C.; Blaschke, T. A multi-scale segmentation/object relationship modelling methodology for landscape analysis. *Ecol. Model.* **2003**, *168*, 233–249. [CrossRef]
54. Arvor, D.; Durieux, L.; Andrés, S.; Laporte, M.A. Advances in geographic object-based image analysis with ontologies: A review of main contributions and limitations from a remote sensing perspective. *ISPRS J. Photogramm.* **2013**, *82*, 125–137. [CrossRef]
55. Davies, Z.G.; Edmondson, J.L.; Heinemeyer, A.; Leake, J.R.; Gaston, K.J. Mapping an urban ecosystem service: Quantifying above-ground carbon storage at a city-wide scale. *J. Appl. Ecol.* **2011**, *48*, 1125–1134. [CrossRef]
56. Al-Kofahi, S.; Steele, C.; VanLeeuwen, D.; Hilaire, R.S. Mapping land cover in urban residential landscapes using very high spatial resolution aerial photographs. *Urban For. Urban Green.* **2012**, *11*, 291–301. [CrossRef]
57. Verbeeck, K.; Van Orshoven, J.; Hermy, M. Measuring extent, location and change of imperviousness in urban domestic gardens in collective housing projects. *Landsc. Urban Plan.* **2011**, *100*, 57–66. [CrossRef]
58. Saltelli, A.; Ratto, M.; Andres, T.; Campolongo, F.; Cariboni, J.; Gatelli, D.; Saisana, M.; Tarantola, S. *Global Sensitivity Analysis: The Primer*; John Wiley & Sons: Hoboken, NJ, USA, 2008.



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).